



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/894,898	06/28/2001	John W. Butzberger	SRI/4438	3387

7590 08/03/2004

THOMASON, MOSER & PATTERSON, LLP  
Attorneys at Law  
SUITE 100  
595 SHREWSBURY AVENUE  
SHREWSBURY, NJ 07702

EXAMINER

PIERRE, MYRIAM

ART UNIT	PAPER NUMBER
----------	--------------

2654

DATE MAILED: 08/03/2004

8

Please find below and/or attached an Office communication concerning this application or proceeding.

# Office Action Summary

Application No.

09/894,898

Applicant(s)

BUTZBERGER ET AL.

Examiner

Myriam Pierre

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE \_\_\_\_ MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☐ Claim(s) \_\_\_\_ is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-36 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

## Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

## Attachment(s)

- ☒ Notice of References Cited (PTO-892)
- ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date 3.4.
- ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_.
- ☐ Notice of Informal Patent Application (PTO-152)
- ☐ Other: \_\_\_\_.

## DETAILED ACTION

### ***Claim Rejections - 35 USC § 112***

1. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

**Claims 27, 29, and 31** are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claims contains “dynamically adjusted threshold”, subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

2. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

**Claims 27, 29, and 31** are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The words “thresholds are dynamically adjusted” are vague and indefinite, because the specification does not describe how this relates is done, and neither do the claims.

***Claim Rejections - 35 USC § 102***

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless --

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

**Claims 1, 2, and 8-35** are rejected under 35 U.S.C. 102(b) as being anticipated by Brown (5,719,997).

**As for claim 1**, Brown teaches a method for allocating memory in a speech recognition system comprising the steps of:

Inherently acquiring a first set of data structures that contain a grammar, a word sub-grammar, a phone sub-grammar and a state sub-grammar, each of the sub-grammars related to the grammar (**Fig 1, col. 3, lines 41-42**);

acquiring a speech signal (**speech input, column 1, lines 26-28**);

performing a probabilistic search using the speech signal as an input, and using the grammar and inherent sub-grammars as possible inputs ("**...mixture probability processor...grammar processor**" column 1, lines 39-40);

and allocating memory for one of the sub-grammars when a transition to that sub-grammar is made during the probabilistic search ("**...evolutional grammar instantiated when needed**" column 8, lines 8-18, lines 11-23 and column 2, lines 16-18; "**de-instantiated...**" column 2, lines 23-25").

As to claim 2, Brown teaches that the probabilistic search is a Viterbi beam search ("**beam**" **searching...**"**Viterbi...**", column 1, lines 41-42).

As to claim 8, 9 and 10, Brown teaches of acquiring a second set of data structures that contain a second grammar, a second word sub-grammar, a second phone sub-grammar, and a second state sub-grammar, each of the second sub-grammars related to the second grammar and replacing the previous one while the speech recognizer is operating (**The data structure for instantiations of HMM are used to allocate memory, which will replace grammar by "de-instantiating grammar" that is no longer needed. De-instantiating grammar includes sub-grammars because non-terminal tables are used to define all "sub-grammars" with they system. And non-terminal tables are an EHMM (de-instantiated or Ephemeral HMM) creation table. Instantiated portions of the grammar are de-instantiated are replaced by others that are instantiated. Instantiations and de-instantiations are done during the speech recognition processing. Column 4, lines 19-24; column 2, lines 23-24; column 12, lines 15-16; and column 12, lines 13-14; column 9, lines 58-60; and column 9, lines 11-16**).

Art Unit: 2654

**As to claim 11**, Brown teaches of a speech recognition system, a method for recognizing speech comprising the steps of:

Inherently acquiring a first set of data structures that contain a grammar, a word sub-grammar, a phone sub-grammar and a state sub-grammar, each of the sub-grammars related to the grammar structures **(The data structure for instantiations of HMM are used to allocate memory, the recognition systems includes “phone, “word” “grammar” and “sub-grammars”. Column 4, lines 19-24; column 11, lines 39-41 and column 3, lines 40-45);**

acquiring a speech signal **(speech input, column 1, lines 26-28);**

performing a probabilistic search using the speech signal as an input, and using the grammar and the sub-grammars as possible inputs **(Fig 1);**

allocating memory for one of the sub-grammars when a transition to that sub-grammar is made during the probabilistic search **(Grammar processor (sub-grammars) causes the word probability processor to instantiate (allocate memory), column 8, lines 11-14).**

computing a probability of a match between the speech signal and an element of the sub-grammar for which memory has been allocated **(“speech input “ is compared to “stored acoustic features representative of words” (examiner is reading this as ‘memory’) contained in a selected grammar, column 1, lines 26-30”).**

**As to claim 12**, Brown teaches that the probabilistic search is a Viterbi beam search (**“beam” searching...“Viterbi...”**, column 1, lines 41-42).

**As to claim 13-15**, Brown teaches of the step of acquiring a second set of data structures that contain a second grammar, a second word sub-grammar, a second phone sub-grammar, and a second state sub-grammar, each of the second sub-grammars related to the second grammar and replacing the previous one while the speech recognizer is operating. **(The data structures for instantiations of HMM are used to allocate memory, which will replace grammar by “de-instantiating grammar” that is no longer needed. De-instantiating grammar includes sub-grammars because non-terminal tables are used to define all “sub-grammars” with they system. And non-terminal tables are an EHMM (de-instantiated or Ephemeral HMM) creation table. Column 4, lines 19-24; column 2, lines 23-24; column 12, lines 15-16; and column 12, lines 13-14; column 9, lines 58-60).**

Art Unit: 2654

**As to claim 18**, Brown teaches a method for recognizing speech comprising the steps of:

inherently acquiring a first set of data structures that contain a top level grammar and a plurality sub-grammars, each of the sub-grammars hierarchically related to the grammar and to each other (**column 3, lines 14-15 and column 8 lines 65-67 and column 9, lines 1-4**) ;

acquiring a speech signal (**speech input, column 1, lines 14-17**);

performing a probabilistic search using the speech signal as an input, and using the top-level grammar and the sub-grammars as possible inputs ("**...mixture probability processor...grammar processor**" **column 1, lines 39-40**);

allocating memory for specific sub-grammars when transitions to those specific sub-grammars are made during the probabilistic search (**Grammar processor ("sub-grammars") causes the word probability processor to "instantiate" (allocate memory), column 8, lines 11-14**); and

computing probabilities of matches between the speech signal and elements of the sub-grammars for which memory has been allocated ("**speech input** " is compared to "**stored acoustic features representative of words**" (examiner is reading this as 'memory') contained in a selected grammar, **column 1, lines 26-30**").



**As to claim 19**, Brown teaches that the inherent top level grammar includes one or more word sub-grammars, the word sub-grammars including words that are related according to word-to-word transition probabilities (**"N-tuple grammar", column 11, line 45.**)

**As to claim 20**, Brown teaches that each word in a word sub-grammar includes one or more phone sub-grammars, the phone sub-grammars including phones that are related according to phone-to-phone transition probabilities (**"Word probability processor 125 contains a) prototypical word models- Illustratively Hidden Markov Models (HMMs)--for the various words that the system of FIG. 1 is capable of recognizing, based on concatenations of phone representations." column 4, lines 14-17)** .

**As to claim 21**, Brown teaches that each phone in a phone sub-grammar includes one or more state sub-grammars, the state sub-grammars including states that are related according to state-to-state transition probabilities (**"Three state...phone representation...each state...phone probability processor generates tri-phone probabilities from component", column 10, lines 58-64.**)

**As to claim 22**, Brown teaches that the probabilities of matches between the speech signal and elements of the sub-grammars for which memory has been allocated is computed using one or more probability distributions associated with each state (**"Hidden Markov Models with multivariate Gaussian distribution" column 10, lines 38-41"**).

**As to claim 23**, Brown teaches that when a word is allocated in memory, an initial phone for the word and an initial state for the initial phone are also allocated in memory (**"stores a lexicon of phonetic word spellings for the vocabulary words which are keyed on the word index. The Phonetic Lexicon table is used to build an internal structure when instantiating an EHMM"**, column 12, lines 26-27).

**As to claim 24**, Brown teaches that one or more subsequent states are allocated in memory until the end of the phone is reached, the allocation based on a transition probability at each state (**"Phonetic table... are loaded into the grammar processor. Column 13, lines 30-31"**).

**As to claim 25**, Brown teaches that one or more subsequent phones are allocated in memory until the end of the word is reached, the allocation based on a transition probability at each phone (**"...input comprises phone scores that were generated by phone probability processor...column 5, lines 25-29 and Fig 2"**).

**As to claim 26**, Brown teaches that when a state probability falls below a state threshold, the state is de-allocated from memory. (**"...drop below it, it can be safely assumed that that portion of the network relates to input that has already been received and processed and it is at that point that the model is de-instantiated. "** column 2, lines 41-43)

**As to claim 28**, Brown teaches that when a phone probability falls below a phone threshold, the phone is de-allocated from memory (“...the HMM are instantiated when needed and de-instantiated when no longer needed, called EHMMs” column 9, lines 57-60 and “HMM have first risen above a predefined threshold and thereafter all drop below it...process...de-instantiated, column 2, lines 40-45”).

**As to claim 30**, Brown teaches that when a word probability falls below a word threshold, the word is de-allocated from memory. (“...drop below it, it can be safely assumed that that portion of the network relates to input that has already been received and processed and it is at that point that the model is de-instantiated. ” column 2, lines 41-43)

**As to claim 32**, Brown teaches that when all the states associated with a phone are de-allocated from memory, the phone is de-allocated from memory (“By de-instantiated we mean that, at a minimum, phone score processing and the propagation of hypothesis scores into such portions of the grammar, e.g., a particular HMM, column 9, lines 12-15 and grammar comprises of words column 11, lines 39-41 and “HMM are instantiated only as needed and de-instantiated when no longer needed, column 9, lines 57-60)”

**As to claim 33**, Brown teaches that when all the phones associated with a word are de-allocated from memory, the word is de-allocated from memory (**"By de-instantiated we mean that, at a minimum, phone score processing and the propagation of hypothesis scores into such portions of the grammar, e.g., a particular HMM, column 9, lines 12-15 and grammar comprises of words column 11, lines 39-41 and "HMM are instantiated only as needed and de-instantiated when no longer needed, column 9, lines 57-60)"**).

**As to claim 34**, Brown teaches of a method for allocating memory in a speech recognition system comprising the steps of:

acquiring a set of data structures that contain a grammar and one or more sub-grammars related to the grammar; (**"In grammar processor, non-terminal grammatical rules are used to dynamically generate finite-state sub-grammars comprising of word..." column 11, lines 39-40)**

acquiring a speech signal (**"...recognizing speech and other inputs...column 1, lines 14-16"**);

performing an inherent probabilistic search using the speech signal as an input, and using the grammar and the sub-grammars as possible inputs (**"...grammar...is instantiated in response to any particular input utterance.." column 8, lines 55-56"**); and

allocating memory for a selected one or more of the sub-grammars when a transition to the selected sub-grammar is made during the probabilistic search

**(“Rather, as processing of input speech begins, grammar processor causes word probability processor to instantiate...initial portion of the grammar”, column 8, lines 14-16).**

**As to claim 35** in a speech recognition system, a method for recognizing speech comprising the steps of:

(a) acquiring a set of data structures that contain a grammar and one or more sub-grammars related to the grammar. (... **grammatical rules... generate finite-state sub-grammars comprising of word...**” column 11, lines 39-40);

(b) receiving spoken input signal (“...**recognizing speech and other inputs...**column 1, lines 14-16”);

(c) inherently, using one or more of the data structures to recognize the spoken input (“**data structure used to process phone scores, column 4, lines 22-23 and “phone representation is a phonetic model of speech signal...” column 4, lines 6-7”** ;

(d) inherently, while the speech recognition system is operating, acquiring a second set of data structures that contain a second grammar and one or more sub-grammars related to the second grammar (“**Fig 14**”); and

(e) repeating steps (b) and (c), using the second set of data structures in step (c). (“**word probability processor contains.....data structure for instantiation of HMM**” column 4, lines 18-23 and Fig 14).

***Claim Rejections - 35 USC § 103***

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 3, 4, 5, 6, 7 and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Brown (5,719,977), as applied to claim 1, in further view of Ehsani et al (2002/0032564).

As to Claim 3, 6, and 7, Brown does not teach that the set of data structures is sent through a communication channel by a remote computer, or selected thereby or that the set of data structures is generated by the speech recognition system using information provided at least in part by a remote computer

Ehsani teaches that the set of data structures for a voice-user interface is sent through a communication channel by a remote computer and that the set of inherent data structures is generated by the speech recognition system using information provided at least in part by a remote computer (**Voice telephony server with speech recognition for remote access of databases via voice commands (page 11, paragraph 0200)**).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time the invention was made to use a set of data structure on a remote computer to extend the capability to access external data bases or control applications or devices, as taught by Ehsani (paragraph 0200).

**As to Claim 4, 5, 16, and 17,** Brown does not teach that the set of data structures is included in code that defines a web page and data structures is inherently associated with one or more web pages.

Ehsani teaches a set of data structures included in code that defines a web page and data structures inherently associated with one or more web pages ("voice page(s)" or "codes" is/are represented by data (data structure) for both structure and content of the Web page, and "enables interaction with the Web page using audio input from speech" page 13, paragraph (0231).

It would have been *prima facie* obvious to one of ordinary skill in the art at the time the invention was made to use a set of data structure that included code that defines web page(s) for the purpose of giving the user more flexibility. One skilled in the art would have been motivated to generate the claimed invention with a reasonable expectation of success.

Art Unit: 2654

**As to Claim 36**, Brown teaches of a speech recognition system, a method for recognizing speech comprising the steps of:

(b) receiving spoken input signal (**speech input, column 1, and lines 14-17**);

(c) using one or more of the data structures to recognize the spoken input (**Data structure is used for memory, which comes from the word probability, the word probability is getting it's data from spoken input. column 1, 24-28 and column 4, lines 19-24**);

(d) while the speech recognition system is operating, acquiring a second set of data structures from the first remote computer or from a second remote computer, the second set of data structures containing a second grammar and one or more sub-grammars related to the second grammar (**While the speech recognition system is operating, the figure shows that it will loop back to the input signal to find the next words until it reaches the end of the sentence, Fig 1**); and

(e) Inherently, repeating steps (b) and (c), using the second set of data structures in step (c). (**Fig 1 and Fig 14**).

Brown does not teach (a) acquiring from a first remote computer a set of data structures that contain a grammar and one or more sub-grammars related to the grammar;



Ehsani teaches (a) acquiring from a first remote computer a set of data structures that contain a grammar and one or more sub-grammars related to the grammar **(Uses an application of recognition “grammars” via “remote” voice control (Page 11, column 0200)...Grammar such as word, phone, and states are used in data structure. Ehsani describes the recognition “grammar”, which uses “phonetic” transcription, “word” sequences, and probability (states) to process the voice commands (Page 11, column 0212);**

It would have been *prima facie* obvious to one of ordinary skill in the art at the time the invention was made to use the already known data structure and combine it with a remote computer for the purpose of enables users to have greater access to information by using a remote computer. One skilled in the art would have been motivated to generate the claimed invention with a reasonable expectation of success.

***Conclusion***

5. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure as follows:

Buhrke et al (6,006,181) teaches continuous speech recognition using three layers for acoustic data, phone data and word sequence.

Bennett et al (6,665,640) teaches an interactive speech-enable dialog system of using HMMs to grammars and dictionaries using SQL databases for query.

Huang et al (5,937,384) teaches using context-dependent Hidden Markov models to represent speech.

Sorsa (6,424,945) teaches downloading state-dependent grammars to a mobile terminal using modest memory.

White et al (6,408,272) teaches downloading grammar.

Knittle (5,758,319) teaches generating a language model dynamically during the application, and then loading the new language model.

Brown (5,907,634) teaches of using speech recognition grammar as non-terminals to represent a finite-state grammar, (EHMM and HMM) in which words are broken down into phrase and phonetic processing.


6. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Myriam Pierre whose telephone number is 703-605-1196. The examiner can normally be reached on Monday – Friday from 5:30 a.m. - 2:00p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on 703-306-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MP

7/20/2004



TĀLIVALDIS IVARS ŠMITS  
PRIMARY EXAMINER